MATHEMATICS

A NOTE ON EIGENVECTORS OF BOUNDED LINEAR OPERATORS

BY

A. J. STAM

(Communicated by Prof. A. C. ZAANEN at the meeting of December 18, 1965)

1. *Introduction*

In sections 1 and 2 of this paper A, B, C, D are bounded linear operators on a Banach space X into X. Elements of X are written $x, y, \ldots$, with or without indices. The spectral radius of A is denoted by ϱ_A.

We consider the equation

$$Ax = \lambda x, \tag{1}$$

where $\lambda = \|A\| \exp(i\varphi)$, φ real. Our main result (theorem 1) will be the following. If $A = B+C$, $\|A\| = \|B\| + \|C\|$, $BC = CB$, and x satisfies (1), then also $Bx = \beta x$, $Cx = \gamma x$, with $\beta = \|B\| \exp(i\varphi)$, $\gamma = \|C\| \exp(i\varphi)$.

If the space of bounded linear operators on X into X is strictly convex, i.e. if $\|A\| = \|B\| + \|C\|$ implies linear dependence of B and C, the assertion is trivial. So to find applications we have to look for examples of not strictly convex Banach algebras. We mention

a. Let X be the space of sequences $u = \{u_n\}$ of complex numbers with $\sum |u_n| < \infty$ and $\|u\| \overset{\text{df}}{=} \sum |u_n|$. The operator A with

$$(Au)_n \overset{\text{df}}{=} \sum_{k=1}^{\infty} u_k\, a_{kn}, \qquad n = 1, 2, \ldots,$$

has norm $\sup\limits_k \sum\limits_n |a_{kn}|$.

So if $b_{kn} \geqq 0$, $c_{kn} \geqq 0$ for all k and n, and $\sum\limits_n b_{kn} = \alpha \sum\limits_n c_{kn}$, $k = 1, 2, \ldots$, we have $\|B+C\| = \|B\| + \|C\|$ in a nontrivial way. Important examples of this type are provided by Markov transition matrices. Some applications are given in section 3.

b. The space of countably additive finite set functions on R_n, with convolution as product and total absolute variation as norm is a not strictly convex Banach algebra, as is seen e.g. by adding two finite measures or two countably additive finite set functions that are restricted to disjoint sets. In section 4 we apply theorem 1 to the integral equation

$$g(x) = \int g(x-\tau)\, dF(\tau)$$

in the unknown Borel function g.

Our statement on (1) refers to the point spectrum of A. The question arises whether similar results hold for the residual or continuous spectrum. Theorem 2 gives a partial answer, referring to the set of λ for which the inverse of $\lambda I - A$, if existent, could not be bounded.

Finally, under an additional condition, theorems 1 and 2 continue to hold with $\|A\|$, $\|B\|$, $\|C\|$ replaced by ϱ_A, ϱ_B, ϱ_C.

2. *Theorems and proofs*

We need the following lemmas:

Lemma 1. *Let* $\delta = \|D\| \exp(-i\psi)$, ψ *real, and* $t = \tau \exp(i\psi)$, $\tau > 0$. *Then for every complex* s

$$\lim_{\tau\to\infty} \|e^{-(s+t)\delta} e^{(s+t)D} - e^{-t\delta} e^{tD}\| = 0. \tag{2}$$

Proof. Making use of the norm inequalities, the relation $t\delta = \tau|\delta|$ and the inequality between arithmetic and quadratic means we find

$$\|e^{-(s+t)\delta} e^{(s+t)D} - e^{t\delta} e^{tD}\| \leqq$$

$$\sum_{k=0}^{\infty} \left| e^{-(s+t)\delta} \frac{(s+t)^k}{k!} - e^{t\delta} \frac{t^k}{k!} \right| \|D^k\| \leqq$$

$$\sum_{k=0}^{\infty} \frac{|\tau\delta|^k}{k!} e^{-|\tau\delta|} \, |e^{-s\delta}(1+s/t)^k - 1| \leqq$$

$$\left\{ \sum_{k=0}^{\infty} \frac{|\tau\delta|^k}{k!} e^{-|\tau\delta|} \, |e^{-s\delta}(1+s/t)^k - 1|^2 \right\}^{\frac{1}{2}} =$$

$$[1 - \exp(-s\delta + s|\tau\delta|/t) - \exp(-\bar{s}\bar{\delta} + \bar{s}|\tau\delta|/\bar{t})$$

$$+ \exp\{-s\delta - \bar{s}\bar{\delta} + |\tau\delta|(s/t + \bar{s}/\bar{t} + |s|^2/|t|^2)\}]^{\frac{1}{2}}$$

$$= \{-1 + \exp(|s^2\delta|\tau^{-1})\}^{\frac{1}{2}}.$$

Lemma 2. *Let* $\delta = \varrho_D \exp(-i\psi)$, ψ *real, and* $t = \tau \exp(i\psi)$, $\tau > 0$. *If there is a constant* M *such that*

$$\|D^n\| \leqq M \varrho^n{}_D, \qquad n = 1, 2, \ldots, \tag{3}$$

then for every complex s

$$\lim_{\tau\to\infty} \|e^{-(s+t)\delta} e^{(s+t)D} - e^{-t\delta} e^{tD}\| = 0. \tag{4}$$

Proof. Lemma 2 is proved in the same way as lemma 1. We have $\|D^k\| \leqq M|\delta|^k$ by (3).

The relation (3) may be formulated as a condition on the rate of convergence of $\|D^n\|^{1/n}$ to ϱ_D, viz.

$$1/n \log \|D^n\| - \log \varrho_D = 0(n^{-1}).$$

By lemma 1 we may prove the following theorems.

Theorem 1. *Let*

$$A = B + C, \tag{5}$$

$$BC = CB, \tag{6}$$

$$\|A\| = \|B\| + \|C\|. \tag{7}$$

Then, if x satisfies the equation

$$Ax = \lambda x, \tag{8}$$

with $\lambda = \|A\| \exp(i\varphi)$, φ real, we must have

$$Bx = \beta x \text{ with } \beta = \|B\| \exp(i\varphi), \tag{9}$$

$$Cx = \gamma x \text{ with } \gamma = \|C\| \exp(i\varphi). \tag{10}$$

Proof. If T is a bounded linear operator on X, the relation $Tx = \vartheta x$ holds if and only if $e^{sT}x = e^{s\vartheta}x$ for every complex s. So from (8)

$$x = e^{-\lambda t} e^{tA} x$$

for every complex t. Therefore, since $BC = CB$,

$$e^{sB}x - e^{s\beta}x = e^{-\lambda t} e^{tC}(e^{(s+t)B} - e^{s\beta} e^{tB})x.$$

Take $t = \tau \exp(-i\varphi)$, $\tau > 0$. Then $\lambda t = \tau\|A\| = \tau\|B\| + \tau\|C\| = \beta t + \|tC\|$, so that

$$e^{sB}x - e^{s\beta}x = e^{s\beta - \|tC\|} e^{tC}(e^{-(s+t)\beta} e^{(s+t)B} - e^{-t\beta} e^{tB})x.$$

Since $\|e^{tC}\| \leqq e^{\|tC\|}$, we see from lemma 1, letting $\tau \to \infty$, that

$$e^{sB}x - e^{s\beta}x = 0.$$

Theorem 2. *If* (5), (6) *and* (7) *hold, and there is a sequence x_n, $n = 1, 2, \ldots$, with $\|x_n\| = 1$ for every n, such that*

$$\lim_{n\to\infty} \|Ax_n - \lambda x_n\| = 0, \tag{11}$$

where $\lambda = \|A\| \exp(i\varphi)$, φ real, then

$$\lim_{n\to\infty} \|Bx_n - \beta x_n\| = 0,$$

$$\lim_{n\to\infty} \|Cx_n - \gamma x_n\| = 0,$$

with $\beta = \|B\| \exp(i\varphi)$, $\gamma = \|C\| \exp(i\varphi)$.

Proof. If T is a bounded linear operator on X, the relation

$$\lim_{n\to\infty} \|Tx_n - \vartheta x_n\| = 0,$$

with $\|x_n\| = 1$, $n = 1, 2, \ldots$, holds if and only if

$$\lim_{n\to\infty} \|e^{sT}x_n - e^{s\vartheta}x_n\| = 0$$

for every complex s. So from (11)

$$x_n = e^{-\lambda t} e^{tA} x_n + y_n(t)$$

with $\lim\limits_{n\to\infty} \|y_n(t)\| = 0$ for every fixed t. So

$$e^{sB} x_n - e^{s\beta} x_n = e^{-\lambda t} e^{tA} (e^{sB} - e^{s\beta}) x_n + (e^{sB} - e^{s\beta}) y_n(t),$$

$$\|e^{sB} x_n - e^{s\beta} x_n\| \leqq \|e^{-\lambda t} e^{tA} (e^{sB} - e^{s\beta})\| + \|e^{sB} - e^{s\beta}\| \, \|y_n(t)\|. \tag{12}$$

If we take $t = \tau \exp(-i\varphi)$, $\tau > 0$, it follows from lemma 1, in the same way as in the proof of theorem 1, that we may choose $\tau = \tau_1$ such that the first term in the right hand side of (12) is bounded by $\frac{1}{2}\varepsilon$. For this fixed τ_1 the second term tends to zero if $n \to \infty$.

The following versions of theorem 1 and 2, with norm replaced by spectral radius, are obtained by means of lemma 2:

Theorem 3. *Let*

$$A = B + C, \tag{13}$$

$$BC = CB, \tag{14}$$

$$\varrho_A = \varrho_B + \varrho_C, \tag{15}$$

$$\|B^n\| \leqq M \varrho^n{}_B, \quad \|C^n\| \leqq N \varrho^n{}_C, \qquad n = 1, 2, \ldots. \tag{16}$$

Then, if x satisfies the equation

$$Ax = \lambda x,$$

with $\lambda = \varrho_A \exp(i\varphi)$, φ real, we have

$$Bx = \beta x, \text{ with } \beta = \varrho_B \exp(i\varphi),$$

$$Cx = \gamma x, \text{ with } \gamma = \varrho_C \exp(i\varphi).$$

Proof. The proof is similar to that of theorem 1, except that now lemma 2 is used and $\exp(-t\varrho_c) \|\exp(tC)\| \leqq N$ by the second part of (16).

Theorem 4. *If* (13), (14), (15) *and* (16) *hold and there is a sequence x_n, $n = 1, 2, \ldots$, with $\|x_n\| = 1$ for every n, such that*

$$\lim_{n\to\infty} \|Ax_n - \lambda x_n\| = 0,$$

with $\lambda = \varrho_A \exp(i\varphi)$, φ real, then

$$\lim_{n\to\infty} \|Bx_n - \beta x_n\| = 0,$$

$$\lim_{n\to\infty} \|Cx_n - \gamma x_n\| = 0,$$

with $\beta = \varrho_B \exp(i\varphi)$, $\gamma = \varrho_C \exp(i\varphi)$.

The proof is similar to that of theorem 2, the same remarks applying as with theorem 3.

3. *Applications to Markov processes.*

Let $P=\{p_{ij},\ i=1,2,\ldots,\ j=1,2,\ldots\}$ be the transition matrix of a Markov chain. The stationary probability distribution $\{\pi_1, \pi_2, \ldots\}$, if existing, of this process satisfies the equation

$$\pi_j = \sum_i \pi_i p_{ij}, \qquad j=1,2,\ldots. \tag{17}$$

If P' and P'' are Markov transition matrices, such that $P'P''=P''P'$ and $P=\varrho P'+\sigma P''$ with $\varrho>0$, $\sigma>0$, $\varrho+\sigma=1$, then by theorem 1 the π_j also satisfy (17) with p_{ij} replaced by p'_{ij} and p_{ij} replaced by p''_{ij}. This is seen by taking the space X as in example (a) of section 1. Similar considerations apply to any eigenvector of P with eigenvalue on the unit circle, that might exist.

If $\{P(t),\ t\geqq 0\}$ is the transition matrix of a continuous parameter Markov process with countable state space, the stationary distribution, if present, satisfies (17) for every $t\geqq 0$, but more important in practice is the stationary equation in terms of the Q-matrix, which is defined by

$$q_{ij} \overset{\mathrm{df}}{=} \lim_{t\to 0} t^{-1}\{p_{ij}(t)-\delta_{ij}\}.$$

(*see* Chung [2], § II. 2, II. 3). We have

$$q_{ii}\leqq 0, \qquad i=1,2,\ldots,\ q_{ij}\geqq 0, \qquad j\neq i, \tag{18}$$

and

$$\sum_{j\neq i} q_{ij}\leqq |q_{ii}|, \qquad i=1,2,\ldots. \tag{19}$$

Here we assume that Q is bounded, i.e.

$$|q_{ii}|\leqq M<\infty, \qquad i=1,2,\ldots. \tag{20}$$

This implies, by (19), that Q has finite norm, and it is sufficient for the relation

$$\sum_j q_{ij}=0, \qquad i=1,2,\ldots \tag{21}$$

to hold. A stationary probability distribution $\{\pi_1, \pi_2, \ldots\}$ of the process, if present, satisfies

$$\sum_i \pi_i q_{ij}=0, \qquad j=1,2,\ldots. \tag{22}$$

Now let Q' and Q'' be matrices satisfying (18) and (21), whereas $Q'Q''=Q''Q'$ and $Q=Q'+Q''$, so that Q' and Q'' have finite norm. Then $\{\pi_1, \pi_2, \ldots\}$ satisfies (22) with q_{ij} replaced by q'_{ij} or by q''_{ij}. This follows from our assertion on (17), since by (18), (20) and (21) there is $\lambda>0$ such that $A=\lambda Q+I$ is a Markov matrix. Then (22) may be written

$$\sum_i \pi_i a_{ij}=\pi_j, \qquad j=1,2,\ldots. \tag{22'}$$

Now $A' = 2\lambda Q' + I$ and $A'' = 2\lambda Q'' + I$ also are Markov matrices, and $A = \frac{1}{2}A' + \frac{1}{2}A''$, $A'A'' = A''A'$, so that

$$\sum_i \pi_i a_{ij}' = \sum_i \pi_i a''_{ij} = \pi_j, \qquad j = 1, 2, \ldots,$$

which proves our assertion.

As more specific examples we give alternative proofs and slight extensions of two theorems on derived Markov chains. Let P be a transition matrix and $a_0, a_1, a_2, \ldots$ be a sequence of nonnegative numbers with sum 1. Then

$$R \overset{\text{df}}{=} \sum_{k=0}^{\infty} a_k P^k \tag{23}$$

again is a Markov transition matrix. It is called the matrix derived from P by the sequence $a_0, a_1, a_2, \ldots$ and may be considered as the transition matrix of the Markov chain arising by sampling a chain with transition matrix P at independent equidistributed intervals. We refer to COHEN [3], [4] and STAM [6].

The period of the sequence $a_0, a_1, a_2, \ldots$ is the largest natural number d for which $\sum_{h=0}^{\infty} a_{hd} = 1$. It was shown in [4] and [6] that, apart from periodicities in the deriving sequence or in the matrices P^n, either P and R have the same stationary distributions or both have none. Clearly every solution of $Px = x$ with $x \in X$, where X and the linear operation Px are as in example a of section 1, also is a solution of $Rx = x$. On the other hand, let $y \in X$ be a solution of $Rx = x$. Then $R^n y = y$, $n = 1, 2, \ldots$. Now

$$R^n = \sum_{k=0}^{\infty} a_k^{(n)} P^k, \tag{24}$$

where $a_0^{(n)}, a_1^{(n)}, a_2^{(n)}, \ldots$ is the n-fold convolution of $a_0, a_1, a_2, \ldots$. If the sequence $a_0, a_1, a_2, \ldots$ has period d, there are n and h such that $a_{hd}^{(n)} > 0$, $a_{hd+d}^{(n)} > 0$. So by theorem 1

$$P^{hd} y = y, \; P^{hd+d} y = y,$$

which implies $P^d y = y$. So the stationary equations for R and P^d have the same set of solutions. Our method also shows that any solution of $Rx = \vartheta x$ with $|\vartheta| = 1$ must satisfy $P^d x = \vartheta^r x$, where r is the least n with $a_{hd}^{(n)} > 0$, $a_{hd+d}^{(n)} > 0$ for some h.

Let C be a closed set of states for P, i.e. $\sum_{j \in C} p_{ij} = 1$, $i \in C$. Then C is also closed for R. For any j in the complement T of C, let η_j be the probability that the chain with transition matrix P, starting in j, will never enter C, and let ξ_j be defined similarly for R. It was shown in [4], that $\eta_j = \xi_j$, $j \in T$.

It is well known (see Feller [5], XV. 8) that $\{\eta_j,\ j \in T\}$ is the maximal solution bounded by 1 of

$$y_j = \sum_{k \in T} p_{jk} y_k, \qquad j \in T.$$

Let Y be the space of bounded sequences $y = \{y_j,\ j \in T\}$ of complex numbers. The operator A on Y with

$$(Ay)_j \overset{\mathrm{df}}{=} \sum_{k \in T} a_{jk} y_k$$

is bounded if $\sup\limits_{j \in T} \sum\limits_{k \in T} |a_{jk}| < \infty$ and then $\|A\| = \sup\limits_{j \in T} \sum\limits_{k \in T} |a_{jk}|$. The equations for $\{\eta_j,\ j \in T\}$ and $\{\xi_j,\ j \in T\}$ may be written $\eta = P_T \eta$, $\xi = R_T \xi$, where $P_T \overset{\mathrm{df}}{=} \{p_{ij},\ i \in T,\ j \in T\}$. We intend to show that the equations $y = R_T y$ and $y = (P^d)_T y$, where d is the period of the sequence $a_0, a_1, a_2, \ldots$, have the same set of solutions in Y. Since the probability of staying infinitely in T is the same for the chains with transition matrices P and P^d, our result implies $\eta_j = \xi_j$, $j \in T$.

From the relation $(P^m)_T = (P_T)^m$, $m = 1, 2, \ldots$, it follows that any solution $y \in Y$ of $y = (P^d)_T y$ satisfies $y = R_T y$. On the other hand, let $z \in Y$ be a non null solution of $y = R_T y$. Then $z = (R_T)^m z$, $m = 1, 2, \ldots$. Since $\|(R_T)^m\| \leqq 1$ and $z \neq 0$, we have $\|(R_T)^m\| = 1$. Now

$$(R_T)^m = \sum_{k=0}^{\infty} a_k^{(m)} (P_T)^k, \qquad m = 1, 2, \ldots, \tag{25}$$

so that $\|(P_T)^k\| = 1$ for every k with $a_k^{(m)} > 0$. Then (7) is satisfied for (25) and we may conclude that $z = (P_T)^d z = (P^d)_T z$ in the same way as before (below (24)).

4. *An integral equation of convolution type.*

As a further application of theorem 1 we consider the integral equation

$$\lambda f(t) = \int_{-\infty}^{+\infty} f(t-\tau)\, dF(\tau), \ -\infty < t < \infty, \tag{26}$$

where the unknown f is a bounded complex valued Borel function and F is a complex valued function of total absolute variation $|\lambda|$. This equation was treated by other techniques by Choquet and Deny [1].

We may rewrite (26) as

$$f(t) = \int_{-\infty}^{+\infty} f(t-\tau) \psi(\tau)\, dH(\tau), \ -\infty < \tau < \infty \tag{26'}$$

with $|\psi(\tau)| = 1$ and H nondecreasing with $H(+\infty) - H(-\infty) = 1$.

To apply theorem 1 we take for X the space of all bounded Borel functions on $(-\infty, +\infty)$ with norm $\|g\| \overset{\mathrm{df}}{=} \sup\limits_t |g(t)|$. The linear operator A on X into X with

$$(Ag)(t) \overset{\mathrm{df}}{=} \int g(t-\tau)\, dG(\tau) \tag{27}$$

is bounded if $\mathrm{Var}\,(G)<\infty$ and then $\|A\|=\mathrm{Var}\,(G)$. Theorem 1 provides the following conclusions on (26):

(a) If $H=\alpha H_1+\beta H_2$, with $\alpha>0$, $\beta>0$, $\alpha+\beta=1$ and H_1, H_2 non-decreasing with total variation 1, then

$$f(t)=\int f(t-\tau)\psi(\tau)dH_k(\tau), \qquad k=1, 2,$$

since operators of the form (27) commute and the norm condition (7) is satisfied.

(b) If $\Delta H(a)\overset{\mathrm{df}}{=}H(a^+)-H(a^-)>0$, we must have

$$f(t)=\psi(a)f(t-a), \ -\infty<t<\infty.$$

This is seen from (a) by taking $H_1(x)$ constant except for a jump in a. For a non null solution to exist it is necessary that

$$\psi(a+b)=\psi(a)\psi(b)$$

for every a, b with $\Delta H(a)>0$, $\Delta H(b)>0$, $\Delta H(a+b)>0$.

(c) If f is continuous and c is a point of increase of H, then

$$f(t)=f(t-c)\gamma(c), \tag{28}$$

where

$$\gamma(c)\overset{\mathrm{df}}{=}\lim_{\varepsilon\to 0^+}\{H(c+\varepsilon)-H(c-\varepsilon)\}^{-1}\int_{c-\varepsilon}^{c+\varepsilon}\psi(\tau)dH(\tau). \tag{29}$$

The relation (28) and the existence of the limit in (29) follow from (a) and the continuity of f by taking for H_1 the restriction of H to $[c-\varepsilon, c+\varepsilon)$ and letting $\varepsilon\to 0$. If H has at least two points of increase whose ratio is irrational, it is easily seen from (28) and the continuity of f, that, if f is not identically zero, we must have $\gamma(c)=\exp\,(i\alpha c)$,

$$f(t)=C\exp\,(i\alpha t), \ -\infty<t<\infty,$$

which on substitution into (26′) implies

$$\psi(\tau)=\exp\,(i\alpha\tau), \text{ a.e.}, \tag{30}$$

with respect to the measure defined by H.

(d) If F or any of its iterated convolutions has an absolutely continuous component, the conclusions of (c) hold, since by (a)

$$f(t)=\int f(t-\tau)h(\tau)d\tau$$

for some $h\in L_1$, which implies that f is continuous.

(e) In the general case no more can be derived from (a) than the following conclusions:

$$\int_E \{f(t)-f(t-\tau)\psi(\tau)\}dH(\tau)=0$$

for every Borel set E, so that $f(t)=\psi(\tau)f(t-\tau)$, a.e. with respect to H for every fixed t.

Now assume that H has at least two points of increase with irrational ratio. Then a necessary condition for (26) to have a solution differing from zero on a set of positive Lebesgue measure, is that (30) holds. This follows from (c), since if f satisfies (26), also Af for any A of the form (27) is a solution, which is continuous if G in (27) is absolutely continuous. If (30) is satisfied, we must have

$$f(t)=C\exp(i\alpha t), \text{ a.e.},$$

with respect to Lebesgue measure. This is seen by considering

$$\varphi(t)\overset{\mathrm{df}}{=}f(t)\exp(-i\alpha t)$$

and applying (c) to $A\varphi$ for every A of the form (27) with G the distribution function of the uniform mass distribution over any Borel set of positive measure.

Institute of Applied Mathematics,
University of Groningen.

REFERENCES

1. Choquet, G. an J. Deny, Sur l'équation de convolution $\mu = \mu * \sigma$. C.R.Ac.Sc. Paris, **250**, 799–801 (1960).
2. Chung, K. L., Markov Chains with Stationary Transition Probabilities. Springer 1960.
3. Cohen, J. W., Derived Markov chains. Indag. Math. **24**, 55–92 (1962).
4. Cohen, J. W., On derived and nonstationary Markov chains. Theory of Probability and its Applications (Moscow) VII, 410–432 (1962).
5. Feller, W., An Introduction to Probability Theory and its Applications I, sec. ed. Wiley, 1957.
6. Stam, A. J., Derived stochastic processes. Compos. Math. **17**, 102–140 (1965).